DSSTox Field Definition File:

National Toxicology Program Bioassay On-line Database (NTPBSI)

Structure-Index Locator File

(last updated 24 April 2008)

Description: Information in this file is intended to provide a minimum level of annotation to the DSSTox SDF (Structure Data Format) file created for the National Toxicology Program (NTP) Bioassay On-line Database (NTPBSI) Chemical Lists. For further explanation of Source-specific fields and background pertaining to the content of this database, a user is encouraged to consult the Source Websites below. Additional information is provided on the DSSTox NTPBSI SDF Download Page http://www.epa.gov/ncct/dsstox/sdf_ntpbsi.html. The DSSTox NTPBSI data file is intended to provide a top-level indexing of the chemical substance and Study Area toxicity data pages contained in the current version of the NTP On-line Database.

** **Note:** The NTPBSI file does not index the entire content of the NTP On-line Database, but only that content related to an explicit chemical exposure. NTP studies for non-chemical stressors such as electromagnetic field studies, or transgenic studies, etc., are not indexed in the DSSTox NTPBSI file but can be located on the NTP Bioassay On-line Database.

Description of ***DSSTox Standard Chemical Fields*** can be found in the Central Field Definition Table located at:
http://www.epa.gov/ncct/dsstox/CentralFieldDef.html

The NTPBSI Structure-Index Locator data file contains no toxicity test data; hence, no ***DSSTox Standard Toxicity Fields*** are employed for this data file. Listed below are ***NTPBSI Source-Specific Fields*** containing chemical substance indexing information (**NTP_CAS_Code**) and URLs (**ChemicalPage_URL**) pointing to the chemical substance-specific data pages in the NTP On-line Database (see Source Websites below). The ***Field Type*** indicates the type of the field, such as numeric, integer, defined text, memo, etc. All ***Units*** and ***Descriptions*** are extracted from Source reference materials unless otherwise noted. ***Allowable Entries*** lists allowed field entries occurring in NTPBSI, separated by slashes for exclusive entries (i.e., cannot occur with another entry) and semicolons or spaces for non-exclusive entries (i.e., can occur with other values). These are defined and explained in the ***Description*** section.

Source Websites: NIEHS's National Toxicology Program (http://ntp.niehs.nih.gov/), NTP On-line Database Main Search Page: http://ntp-apps.niehs.nih.gov/ntp_tox/index.cfm.

Source Contacts: National Institute of Environmental Health Sciences, National Toxicology Program: Beth Bowden, email: bowden1@niehs.nih.gov or Joe Roycroft, email: roycroft@niehs.nih.gov

SDF Usage Notes:
Each DSSTox SDF file contains a single **STRUCTURE** field. For each chemical record, the **STRUCTURE** field entry directly corresponds to the content of the **STRUCTURE_...** fields. The **STRUCTURE_Shown** field documents the relationship between what is displayed in the **STRUCTURE** field and the actual tested chemical substance, i.e. **TestSubstance_...** fields, with the latter corresponding directly to the toxicity data field entries. Commercial chemical relational database (CRD) applications may automatically insert one or more structure identifier fields upon import or export of an SDF file (e.g., Formula, FW or Mol_ID), fields that may augment or duplicate one or more of the DSSTox Standard Chemical Fields. Users are cautioned that fields containing null values in the first record of the SDF will be reordered upon import into most applications; for this reason, the word "blank" has been inserted into null fields in Record 1 of DSSTox SDF files and can be deleted after SDF import. Users are additionally cautioned that some fields (**STRUCTURE_SMILES** and **STRUCTURE_InChI**, in particular) may exceed the 200 character limit specified in the MDL CTFiles SDF standard (see http://www.epa.gov/ncct/dsstox/MoreonSDF.html), and that some CRD applications may insert a line break or truncate these fields upon SDF import or export. Finally, CRD application-specific molecular header information in the SDF file is deleted in the final DSSTox SDF files; users running CRD applications requiring a unique molecule header upon import of the SDF can specify either **DSSTox_RID** or the **DSSTox_FileID** be used. Upon SDF import, **DSSTox_CID** can be used to identify and manage chemical structure duplicates and **DSSTox_Generic_SID** can be used to identify common Test Substances across and within DSSTox files (similar to CASRN-substance, but available for all DSSTox substances and further distinguishes among different purity/grade substances).

As an MS Word document, the following table is best viewed onscreen using either Normal or Web Layout View in Landscape page orientation.

Field Name	***Field Type***	***Units***	***Allowable Entries***	***Description***	***Comments***
			NTPBSI Source-Specific Fields		
NTP_StudyArea_ ImmunoTox *(no spaces)*	integer		0/ 1/	One of the 4 Study Areas (Immunology) for which bioassay data may be available, indexed by chemical substance, in the NTP Bioassay On-Line Database of the NIEHS National Toxicology Program. Indicator values indicate the presence or absence of study data in this Study Area. 1 = study data available; 0 = no study data available	

NTP_StudyArea_ GeneTox ***(no spaces)***	integer		0/ 1/	One of the 4 Study Areas (Genetox) for which bioassay data may be available, indexed by chemical substance, in the NTP Bioassay On-Line Database of the NIEHS National Toxicology Program. Indicator values indicate the presence or absence of study data in this Study Area. 1 = study data available; 0 = no study data available	
NTP_StudyArea_ CancerChronicTox ***(no spaces)***	integer		0/ 1/	One of the 4 Study Areas (Bioassay) for which bioassay data may be available, indexed by chemical substance, in the NTP Bioassay On-Line Database of the NIEHS National Toxicology Program. Indicator values indicate the presence or absence of study data in this Study Area. 1 = study data available; 0 = no study data available	
NTP_StudyArea_ DevelopTox ***(no spaces)***	integer		0/ 1/	One of the 4 Study Areas (Developmental) for which bioassay data may be available, indexed by chemical substance, in the NTP Bioassay On-Line Database of the NIEHS National Toxicology Program. Indicator values indicate the presence or absence of study data in this Study Area. 1 = study data available; 0 = no study data available	
NTP_CAS_Code	defined text		*URL*	Code used by the NTP to index all test chemical substances contained in the on-line NTP database, consists either of the CAS registry number, if available, that is used by the NTP or a character name abbreviation in cases of mixtures, non-chemical stressors, or other cases where CAS registry numbers are unavailable (e.g., GLYCINEBENZA or EMTDP-91). Code is incorporated into the URL of the NTP main chemical substance study page. In a few cases where an updated CAS is available, the DSSTox **TestSubstance_CASRN** will differ from the **NTP_CAS_Code**; a note to this effect will be included in the **Note_NTPBSI** field.	*Field added to NTPBSI_v2a.*
Note_NTPBSI	memo		*URL*	Field used to provide supplementary Source-specific information pertaining to the chemical and toxicity fields.	Note documents few instances where **TestSubstance_CASRN** differs from NTP CAS used in the **NTP_CAS_Code**. Note also documents if record has been

					added to v2a, v3a, etc. or if Study Areas were modified. *Replaces* **ToxicityNote** *field (June 2007).*
ChemicalPages_ URL	memo		*URL*	Internet URL website address for chemical-specific data or content. URL was checked at time of DSSTox data file publication. Please send DSSTox Error Report if website URL address no longer works or is changed.	URLs point to chemical substance-specific toxicity test data pages in the current NTP On-line Database (http://ntp-apps.niehs.nih.gov/ntp_tox/index.cfm). *URLs were previously included in general field Website_URL; new field added (April 2008).*